



Pitch Estimation and Tracking with Harmonic Emphasis On The Acoustic Spectrum

Karimian-Azari, Sam; Mohammadiha, Nasser; Jensen, Jesper Rindom; Christensen, Mads Græsbøll

Published in:

I E E E International Conference on Acoustics, Speech and Signal Processing. Proceedings

DOI (link to publication from Publisher):

[10.1109/ICASSP.2015.7178788](https://doi.org/10.1109/ICASSP.2015.7178788)

Publication date:

2015

Document Version

Early version, also known as pre-print

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Karimian-Azari, S., Mohammadiha, N., Jensen, J. R., & Christensen, M. G. (2015). Pitch Estimation and Tracking with Harmonic Emphasis On The Acoustic Spectrum. *I E E E International Conference on Acoustics, Speech and Signal Processing. Proceedings*, 4330-4334. <https://doi.org/10.1109/ICASSP.2015.7178788>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Pitch Estimation and Tracking with Harmonic Emphasis on the Acoustic Spectrum

April 23, 2015

Sam Karimian-Azari¹, Nasser Mohammadiha²,
Jesper R. Jensen¹, and Mads G. Christensen¹

ska@create.aau.dk

¹Audio Analysis Lab, AD:MT, Aalborg University

²University of Oldenburg



AALBORG UNIVERSITY
DENMARK



Agenda



- ▶ Introduction
 - ▶ Noisy Harmonic Signal Approximation
 - ▶ ML Pitch Estimate from UFE
- ▶ Bayesian Methods
 - ▶ Motivation
 - ▶ HMM
 - ▶ Kalman Filter
- ▶ Numerical Results
- ▶ Conclusion

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space:

HMM

2- Continuous state-space:

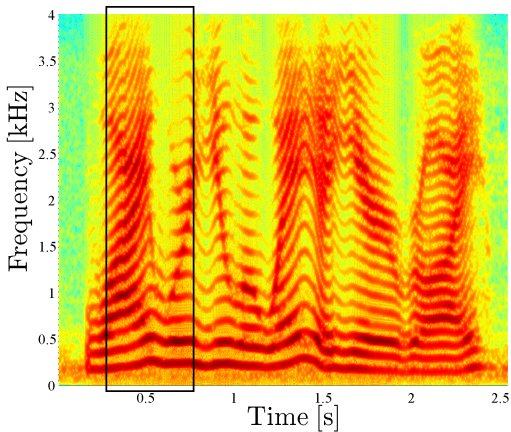
Kalman Filter

Numerical Results

Conclusion

Introduction

Harmonic Signal Model



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

2

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space:
HMM

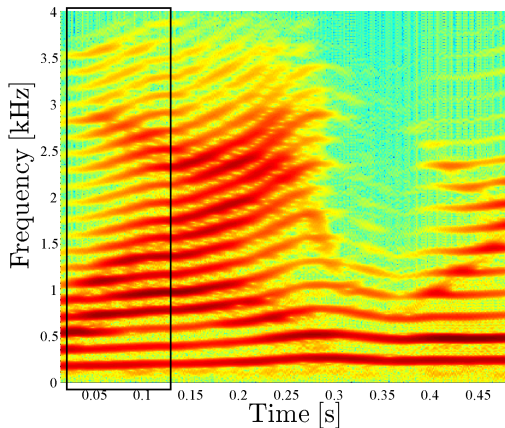
2- Continuous state-space:
Kalman Filter

Numerical Results

Conclusion

Introduction

Harmonic Signal Model



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

3

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space:
HMM

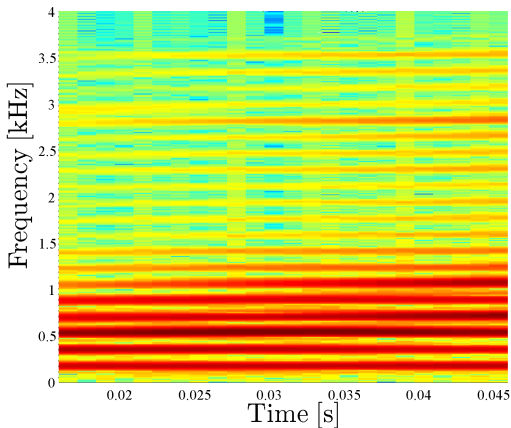
2- Continuous state-space:
Kalman Filter

Numerical Results

Conclusion

Introduction

Harmonic Signal Model



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

4

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space:
HMM

2- Continuous state-space:
Kalman Filter

Numerical Results

Conclusion

Harmonic Signal Model:

$$s(n) = \sum_{l=1}^{L(n)} \alpha_l e^{j(\omega_l(n)n + \varphi_l)},$$

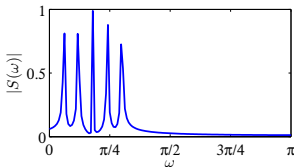
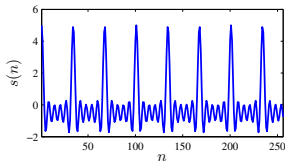
where $\omega_l(n) = l\omega_0(n)$ for $l = 1, \dots, L(n)$,

$L(n)$: number of sinusoids

α_l : real magnitudes

ω_0 : fundamental frequency

φ_l : phases of harmonics



(1)

5

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

Signal Model

Additive noise



The observed signal can be written as a sum of a desired signal $s(n)$ and a noise signal $v(n)$, i.e.,

$$\begin{aligned} x(n) &= s(n) + v(n) \\ &= \sum_{l=1}^L \alpha_l e^{j(\omega_l n + \varphi_l)} + v(n). \end{aligned} \tag{2}$$

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

6

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

At a high narrowband SNR, the harmonic frequency ω_l is perturbed with a real-valued phase-noise [S.Tretter 1985], which has a normal distribution with zero mean and the variance

$$E\{\Delta\omega_l^2(n)\} = \frac{\sigma^2}{2\alpha_l^2} \quad (3)$$

We can approximate $x(n) = \sum_{l=1}^L \alpha_l e^{j(\omega_l n + \varphi_l)} + v(n)$ like

$$x(n) \approx \sum_{l=1}^L \alpha_l e^{j(\omega_l n + \Delta\omega_l(n) + \varphi_l)} \quad (4)$$

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

7

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

Unconstrained frequency estimates (UFE) of the constrained frequencies:

$$\hat{\Omega}(n) = [\hat{\omega}_1(n), \hat{\omega}_2(n), \dots, \hat{\omega}_L(n)]^T \quad (5)$$

$$= \mathbf{d}_L(n) \omega_0(n) + \Delta\Omega(n), \quad (6)$$

where

$$\mathbf{d}_L(n) = [1, 2, \dots, L(n)]^T \quad (7)$$

$$\Delta\Omega(n) = [\Delta\omega_1(n), \Delta\omega_2(n), \dots, \Delta\omega_L(n)]^T, \quad (8)$$

and

$$\begin{aligned} \mathbf{R}_{\Delta\Omega}(n) &= \mathbf{E}\{\Delta\Omega(n)\Delta\Omega^T(n)\} \\ &= \frac{\sigma^2}{2} \text{diag}\left\{\frac{1}{\alpha_1^2}, \frac{1}{\alpha_2^2}, \dots, \frac{1}{\alpha_L^2}\right\}. \end{aligned} \quad (9)$$

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

8

19

Max. Likelihood (ML) Pitch Estimator



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space:
HMM

2- Continuous state-space:
Kalman Filter

Numerical Results

Conclusion

For the time-frame $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-M-1)]^T$, the PDF of the UFE is

$$P(\hat{\Omega}(n)|\omega_0(n)) \sim \mathcal{N}(\mathbf{d}_L(n)\omega_0(n), \mathbf{R}_{\Delta\Omega}(n)). \quad (10)$$

The ML pitch estimator:

$$\hat{\omega}_0(n) = \arg \max_{\omega_0(n)} \log P(\hat{\Omega}(n)|\omega_0(n)) \quad (11)$$

$$= [\mathbf{d}_L^T(n)\mathbf{R}_{\Delta\Omega}^{-1}(n)\mathbf{d}_L(n)]^{-1} \mathbf{d}_L^T(n)\mathbf{R}_{\Delta\Omega}^{-1}(n)\hat{\Omega}(n) \quad (12)$$

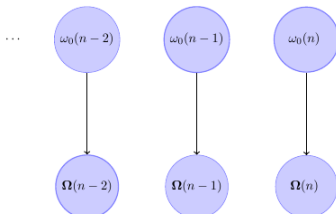
9

19

Bayesian Pitch Estimator

Motivation

- ▶ The ML Estimators are statistically efficient, e.g., the non-linear least-squares (NLS), and the weighted least squares (WLS) [H.Li, et al. 2000], but the minimum variance is limited by the number of samples.
- ▶ Consecutive pitch values are estimated independently



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

10

19

Bayesian Pitch Estimator

Motivation



- Pitch values are usually correlated in a sequence, i.e.,

$$P(\omega_0(n) | \omega_0(n-1), \omega_0(n-2), \dots), \quad (13)$$

that motivate Bayesian methods to minimize an error incorporating prior distributions.

- State-of-the-art methods mostly track pitch estimates in a sequential process without concerning noise statistics.



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

11 Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

Bayesian Pitch Estimator

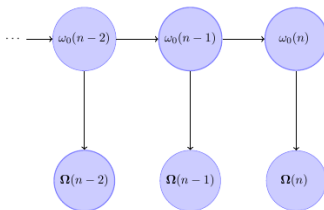
Hypothesis



- 1- Jointly estimate and track pitch incorporating both the harmonic constraints and noise characteristics.
- 2- Estimate the state $\omega_0(n)$ through a series of noisy observations:

$$P(\omega_0(n) | \hat{\Omega}(n), \hat{\Omega}(n-1), \dots) \quad (14)$$

- 3- Recursively update the prior distribution of the pitch value.



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

12

Motivation-Hypothesis

1- Discrete state-space:

HMM

2- Continuous state-space:

Kalman Filter

Numerical Results

Conclusion

Bayesian Pitch Estimator

Discrete state-space (HMM)



Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

13

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

$\omega_0(n)$: Discrete random variable (Hidden states)

$P(\omega_0(n)|\omega_0(n-1))$: Transition probability in a 1st-order Markov model,

$$\text{i.e., } \sum_{\omega_0(n)} P(\omega_0(n)|\omega_0(n-1)) = 1$$

$$\hat{\omega}_0(n) = \arg \max_{\omega_0(n)} \log P(\omega_0(n)|\hat{\Omega}(n), \hat{\Omega}(n-1), \dots) \quad (15)$$

$$= \arg \max_{\omega_0(n)} \log P(\hat{\Omega}(n)|\omega_0(n)) + \log P(\omega_0(n)|\hat{\Omega}(n-1), \dots).$$

The priori distribution is defined recursively like

$$P(\omega_0(n)|\hat{\Omega}(n-1), \hat{\Omega}(n-2), \dots) = \sum_{\omega_0(n-1)} P(\omega_0(n)|\omega_0(n-1)) P(\omega_0(n-1)|\hat{\Omega}(n-1), \dots), \quad (16)$$

where $P(\omega_0(n-1)|\hat{\Omega}(n-1), \dots)$ is the past estimate.

19

Bayesian Pitch Estimator

state-space representation of the pitch continuity



Continuous state-space:

$$\omega_0(n) = \omega_0(n-1) + \delta(n)$$

$$\hat{\Omega}(n) = \mathbf{d}_L(n) \omega_0(n) + \Delta\Omega(n),$$

where $\delta(n) \sim \mathcal{N}(0, \sigma_t^2)$ and $\Delta\Omega(n) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\Delta\Omega}(n))$ are the state evolution and observation noise, respectively.

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

14

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

Bayesian Pitch Estimator

Continuous state-space (Kalman filter)



First, a pitch estimate is predicted using the past estimates as

$$\hat{\omega}_0(n|n-1) = \hat{\omega}_0(n-1|n-1) \quad (17)$$

with the variance

$$\sigma_K^2(n|n-1) = \sigma_K^2(n-1|n-1) + \sigma_t^2. \quad (18)$$

Second, the pitch estimate is updated with the error of

$$\mathbf{e}(n) = \hat{\Omega}(n) - \mathbf{d}_L(n) \hat{\omega}_0(n|n-1). \quad (19)$$

Then, the predicted estimate is updated:

$$\hat{\omega}_0(n|n) = \hat{\omega}_0(n|n-1) + \mathbf{h}_K(n) \mathbf{e}(n) \quad (20)$$

$$\mathbf{h}_K(n) = \sigma_K^2(n|n-1) \mathbf{d}_L^T(n) \left[\mathbf{\Pi}_L(n) \sigma_K^2(n|n-1) + \mathbf{R}_{\Delta\Omega}(n) \right]^{-1}, \quad (21)$$

where $\mathbf{\Pi}_L(n) = \mathbf{d}_L(n) \mathbf{d}_L^T(n)$, and update

$$\sigma_K^2(n|n) = \left[1 - \mathbf{h}_K(n) \mathbf{d}_L(n) \right] \sigma_K^2(n|n-1). \quad (22)$$

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

Conclusion

15

The ML estimator of the covariance matrix among N estimates:

$$\begin{aligned}\mathbf{R}_{\Delta\Omega}(n) &= \mathbf{E}\{\Delta\Omega(n)\Delta\Omega^T(n)\} \\ &= \frac{1}{N} \sum_{i=n-N+1}^n \Delta\Omega(i)\Delta\Omega^T(i),\end{aligned}\tag{23}$$

where $\Delta\Omega(n) = \hat{\Omega}(n) - \hat{\mu}(n)$, and $\mu(n) = \mathbf{E}\{\hat{\Omega}(n)\}$.

Exponential moving average:

$$\hat{\mu}(n) = \lambda \hat{\Omega}(n) + (1 - \lambda) \hat{\mu}(n-1)\tag{24}$$

The forgetting factor $0 < \lambda < 1$ recursively updates the time-varying mean value.

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

16

2- Continuous state-space: Kalman Filter

Numerical Results

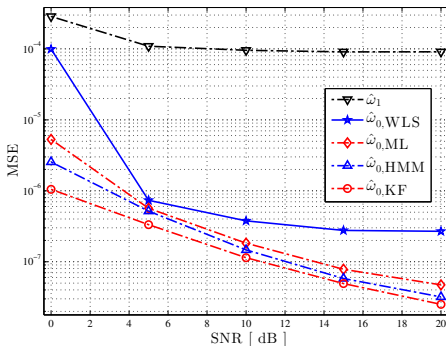
Conclusion

Numerical Results

Synthetic signal



A linear chirp signal ($r = 100$ Hz/s) with $L = 5$ harmonics, random phases, and identical amplitudes during 0.1 s.



$M = 80$, $\omega_0(1) = 400\pi / f_s$, $f_s = 8.0$ kHz, $\sigma_t = \sqrt{2}\pi r / f_s^2$, and for the HMM-based pitch estimator, the frequency range $\omega \in [150, 280] \times (2\pi / f_s)$ was discretized into $N_d = 1000$ samples.

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

17 Numerical Results

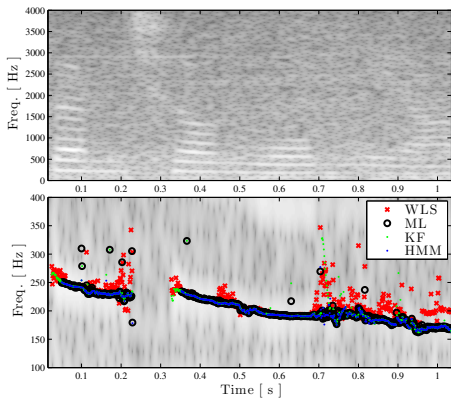
Conclusion

Numerical Results

Real signal



Speech signal + Car noise at SNR= 5 dB.



The MAP order estimation [Djuric 1998], $M = 240$, $\lambda = 0.9$, and $N = 150$.

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space:
HMM

2- Continuous state-space:
Kalman Filter

18 Numerical Results

Conclusion

19

Audio Analysis Lab, AD:MT,
Aalborg University, Denmark

Conclusion



- ▶ For pitch estimation, we have formulated the ML estimate from the UFE.
- ▶ For pitch estimation and tracking, we have proposed HMM- and KF-based methods.
- ▶ Experimental results showed that both HMM- and KF-based methods outperform the corresponding ML pitch estimators.
- ▶ The KF-based method statistically performs better than the HMM-based method, while it tracks pitch changes more accurately than the KF-based method.

Pitch Estimation and Tracking

Sam Karimian-Azari et al.

Introduction

Harmonic Signal Model

Noisy Signal Approx.

ML Pitch Estimation

Proposed Method

Motivation-Hypothesis

1- Discrete state-space: HMM

2- Continuous state-space: Kalman Filter

Numerical Results

19 Conclusion

Thank you!



AALBORG UNIVERSITY
DENMARK

